

Causally nonseparable processes admitting a causal model

Adrien Feix,^{1,2} Mateus Araújo,^{1,2} and Časlav Brukner^{1,2}

¹*Faculty of Physics, University of Vienna, Boltzmannngasse 5, 1090 Vienna, Austria*

²*Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmannngasse 3, 1090 Vienna, Austria*

(Dated: August 23, 2016)

A recent framework of quantum theory with no global causal order predicts the existence of “causally nonseparable” processes. Some of these processes produce correlations incompatible with any causal order (they violate so-called “causal inequalities” analogous to *Bell inequalities*) while others do not (they admit a “causal model” analogous to a *local model*). Here we show for the first time that bipartite causally nonseparable processes with a causal model exist, and give evidence that they have no clear physical interpretation. We also provide an algorithm to generate processes of this kind and show that they have nonzero measure in the set of all processes. We demonstrate the existence of processes which stop violating causal inequalities but are still causally nonseparable when mixed with a certain amount of “white noise”. This is reminiscent of the behavior of Werner states in the context of entanglement and nonlocality. Finally, we provide numerical evidence for the existence of causally nonseparable processes which have a causal model even when extended with an entangled state shared among the parties.

I. INTRODUCTION

It is well-known that quantum mechanics is at odds with naive notions of reality and locality as predicted by Bell’s celebrated theorem [1, 2]. One might wonder whether peculiar quantum features could challenge other fundamental notions, like the concept of *causality*, as well. The *process matrix formalism* of Oreshkov, Costa and Brukner [3] was developed to explore this question—studying the most general causal structures compatible with local quantum mechanics for two parties A and B .

Surprisingly, the formalism predicts causal structures which are “causally nonseparable”: they correspond neither to A being before B nor to B being before A , nor to a probabilistic mixture thereof. These causal structures can produce correlations incompatible with any definite causal order, violating so-called “causal inequalities” [3, 4].¹

However, the empirical relevance of these results is still completely unclear. Do they appear in some physical situations or are they merely a mathematical artifact of the process matrix formalism?

For three and more parties, there are causally nonseparable processes whose physical realization is known—one instance is the “quantum switch” [6] where the causal order between two parties A and B is controlled by a quantum system belonging to a third party C . Processes of this kind, however, *cannot violate* causal inequalities [7, 8]: they admit a “causal model”, i.e., a causally separable process is capable of reproducing their correlations. Their causal nonseparability can only be certified through device-dependent “causal tomography” or “causal witnesses” [7]. This is analogous to states

which are entangled but cannot violate Bell inequalities, i.e., for which a “local model” exists [9].

Since the only causally nonseparable processes known to be physically implementable have a causal model, it is tempting to conjecture that the inability to violate causal inequalities without [10] or with [8] operations extended to shared entangled states by all parties *singles out* the physical causal structures from unphysical ones. Ref. [8] contains an example of a tripartite process matrix with a causal model but which does not remain causal under extensions, i.e., is not “extensibly causal”, demonstrating the difference between the two notions for more than two parties.

In this paper, we provide an example of a bipartite causally nonseparable process with a causal model. Furthermore, we give numerical evidence that *bipartite* nonseparable processes exist which do not violate causal inequalities, even when extended with entanglement. No physical interpretation is known for these processes.

The paper is organized as follows: Sec. II introduces the process matrix formalism and the definitions of causal nonseparability and causal inequalities. In Sec. III, we define a class of two-party causally nonseparable processes and construct a causal model for them. This shows that the sets of causally nonseparable and causal inequality violating processes are distinct also in the bipartite case. Since the causally nonseparable processes with a causal model can be interpreted as the mixture of physically implementable process with an unphysical process, this gives evidence that they are not implementable in nature.

In Sec. IV, we provide an algorithm to construct nonseparable processes with a causal model by composing a random causally separable process with a non-completely positive map on one party. Using a random sample generated by a “hit-and-run” Markov chain [11, 12], we also show that nonseparable processes with a causal model have nonzero measure in the space of all processes.

In Sec. V, we construct a family of “Werner processes”,

¹ For two parties, the operations implemented to violate causal inequalities have to be quantum [3]; surprisingly, for three or more parties, classical operations are sufficient [5].

which mimic the behavior of Werner states [9] with respect to nonlocality and entanglement. The Werner processes’ causal nonseparability is more resistant to the introduction of “white noise” than its ability to violate causal inequalities. This shows that the analogy between causal nonseparability and causal inequalities on the one hand, and entanglement and Bell inequalities on the other, extends beyond what was previously known [3, 7, 8, 13].

Finally, in Sec. VI, we examine the behavior of the processes with added shared entanglement between the parties. While some of the processes we constructed *do violate causal inequalities* when extended in this way (are not “extensibly causal”), numerical calculations indicate that the ability to violate causal inequalities disappears when adding a little white noise, at which the causal nonseparability is preserved. We conjecture that there are processes which are extensibly causal, and yet not physically implementable.

II. CAUSAL NONSEPARABILITY AND CAUSAL INEQUALITIES

Quantum circuits can be thought of as a formalization of causal structures with a definite causal order. They consist of *wires*, representing quantum systems, which connect boxes, representing *quantum operations*. While for quantum circuits, the order of the operations is fixed [14], situations where the order of operations is not well-defined are readily represented in the *process matrix formalism* [3], which can be thought of as a generalization of the quantum circuit formalism. We will briefly introduce the main elements of the formalism; a more detailed introduction to it can be found in Ref. [7].

A quantum operation maps a density matrix $\rho_{A_I} \in A_I$ to a density matrix $\rho_{A_O} \in A_O$ (where A_I (A_O) denotes the space of linear operators on the Hilbert space \mathcal{H}^{A_I} (\mathcal{H}^{A_O})). The most general operations within the quantum formalism are *completely positive* (CP) maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi-Jamiołkowski [15, 16] (CJ) isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A := [(\mathcal{I} \otimes \mathcal{M}_A)(|I\rangle\langle I|)]^T \in A_I \otimes A_O, \quad (1)$$

where \mathcal{I} is the identity map and $|I\rangle := \sum_{j=1}^{d_{\mathcal{H}_I}} |jj\rangle \in \mathcal{H}_I \otimes \mathcal{H}_I$ is a non-normalized maximally entangled state; T denotes matrix transposition in the computational basis.

The CJ-isomorphism can also be used to represent “superoperators” or “processes” which map *quantum maps* to quantum maps, quantum states or probabilities [3, 14, 17–19]. In this paper, we will focus on processes mapping *two quantum operations* ξ_x^a and η_y^b —corresponding to the Choi-Jamiołkowski representation of Alice’s and Bob’s CP maps—to a *probability* (see

Fig. 1). Requiring linearity of probabilities in the operations, we can represent it as

$$p(\xi_x^a, \eta_y^b) := \text{tr}[W \cdot \xi_x^a \otimes \eta_y^b], \quad (2)$$

$$W \in A_I \otimes A_O \otimes B_I \otimes B_O. \quad (3)$$

To ensure positivity of probabilities for all pairs of possible CP maps (as well as for extended operations where the parties share additional entanglement) the “process matrix” W has to be positive semidefinite $W \geq 0$ [3]. The normalization of probabilities implies that $\text{tr}[W \cdot \xi^{\text{CPTP}} \otimes \eta^{\text{CPTP}}] = 1$ for all CJ-representations of completely positive *trace-preserving maps* ξ^{CPTP} and η^{CPTP} [3].

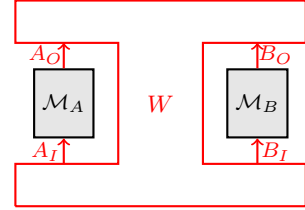


FIG. 1. Representation of a bipartite process W , which linearly maps Alice’s and Bob’s CP maps $\mathcal{M}_A, \mathcal{M}_B$ to a probability. A_I (B_I) represents Alice’s (Bob’s) input Hilbert space and A_O (B_O) Alice’s (Bob’s) output Hilbert space.

We call a process $W^{A \prec B}$ ($W^{B \prec A}$) “causally ordered” if it does not allow for signalling from Bob to Alice (Alice to Bob), which is equivalent to the conditions [7]

$$W^{A \prec B} = \text{tr}_{B_O}[W^{A \prec B}] \otimes \mathbb{1}^{B_O}/d_{B_O}, \quad (4)$$

$$W^{B \prec A} = \text{tr}_{A_O}[W^{B \prec A}] \otimes \mathbb{1}^{A_O}/d_{A_O}. \quad (5)$$

A process matrix $W^{\text{sep}} \in \mathcal{W}_{\text{sep}}$ that can be decomposed into a convex combination ($0 \leq q \leq 1$) of causally ordered processes is “causally separable”:

$$W_{\text{sep}} = qW^{A \prec B} + (1-q)W^{B \prec A}. \quad (6)$$

It was recently shown that one can *efficiently* determine whether a process is causally (non)separable using a *semidefinite program* (SDP). Here we will use the SDP for “random robustness” [7]

$$\begin{aligned} & \min \lambda \\ & \text{s.t. } W = W^{A \prec B} + W^{B \prec A} - \lambda \mathbb{1}^\circ, \\ & W^{A \prec B} = \text{tr}_{B_O}[W^{A \prec B}] \otimes \mathbb{1}^{B_O}/d_{B_O}, \\ & W^{B \prec A} = \text{tr}_{A_O}[W^{B \prec A}] \otimes \mathbb{1}^{A_O}/d_{A_O}, \end{aligned} \quad (7)$$

where $\mathbb{1}^\circ := \mathbb{1}^{A_I A_O B_I B_O} / (d_{A_I} d_{B_I})$. The random robustness $R_r(W)$ is defined as the result of the optimization $R_r(W) := \lambda_{\text{opt}}$.

If $R_r(W) \leq 0$, the SDP gives an explicit decomposition of W into $W^{A \prec B}$ and $W^{B \prec A}$; if $R_r(W) > 0$, the process is not causally separable. The value of $R_r(W)$ is also an operational measure of “causal nonseparability”. It is related to the minimal amount of “white noise” $\mathbb{1}^\circ$ that

needs to be mixed with the process to make it causally separable. That is, for $\gamma \geq R_r(W)$, the process $(\gamma \mathbb{1}^\circ + W)/(1 + \gamma)$ is causally separable.

A so-called “dual SDP” to (7) can then provide the optimal “causal witness”, i.e., a hermitian operator S such that $\text{tr}[W_{\text{sep}} S] \geq 0$ for all causally separable processes W_{sep} [7]. The property $\text{tr}[W S] < 0$ can be verified experimentally by measuring a set of operators for Alice and Bob, certifying that the process W is not causally separable. Note that, in analogy to entanglement witnesses [20], this certification of causal nonseparability relies on a partial tomography of the process and thus requires trust in Alice’s and Bob’s local operations: it is “device-dependent”.

It is well-known that the entanglement of a quantum state can be certified *device-independently* (without requiring trust Alice’s and Bob’s operations) if the probability distribution resulting from a set of measurements violates a Bell inequality [21]. In an analogous way, causal nonseparability can be device-independently confirmed using *causal inequalities* [3–5, 7, 8], where the “non-causal” correlations between Alice and Bob alone suffice to show that the process they use is not causally separable, without additional trust in their local operations.

The condition for a probability distribution to be “causal”, i.e., *not* to violate any causal inequality, is simply that it can be decomposed into a convex combination of a probability distribution which is no-signaling from Bob to Alice² ($p_{A \prec B}$) and a probability distribution which is no-signaling from Alice to Bob ($p_{B \prec A}$) [4]:

$$p_{\text{causal}} = q p_{A \prec B} + (1 - q) p_{B \prec A}. \quad (8)$$

Note that the correlations generated by a causally ordered process cannot violate any causal inequality.

For the scenario where Alice (Bob) has an input bit x (y) and outputs one bit a (b), one causal inequality is a bound on the probability of success of the “guess your neighbor’s input” (GYNI) [4]:

$$p_{\text{GYNI}} := \frac{1}{4} \sum_{x,y} p(a = y, b = x | x, y) \leq \frac{1}{2}. \quad (9)$$

Some valid processes and local strategies which result in correlations violating (9) are described in Ref. [4].

The relation between causally nonseparable processes and the violation of causal inequalities is not yet fully understood. On the one hand, there exist causally nonseparable processes that can be physically implemented [22] but have a *causal model*—they do not violate causal inequalities [7, 8]. An example of such a process is the

“quantum switch” [6]. On the other hand, there are processes that can violate causal inequalities, but it is not known if they can be realized in nature, prompting the conjecture that only processes with a causal model are physically implementable [10].

Another natural feature to investigate is the (in)ability for a process to violate causal inequalities, even when extended with an entangled state shared by all parties. We will call processes which do not allow for such a violation “*extensibly causal*” [8] and come back to this concept in detail in Sec. VI.

In the bipartite case, all previously known nonseparable processes violate causal inequalities and it is not clear if nonseparable processes with a causal model even exist [4]. In the next section, we explicitly provide a class of nonseparable bipartite processes that allow for a causal model.

III. CAUSALLY NONSEPARABLE PROCESSES WITH A CAUSAL MODEL

We will consider the following class of processes (all the operators are understood to act on qubits, $d_{A_I} = d_{A_O} = d_{B_I} = d_{B_O} = 2$):

$$\begin{aligned} W^{A \prec B} &:= \mathbb{1}^\circ + \frac{1}{12}(\mathbb{1} Z Z \mathbb{1} + \mathbb{1} X X \mathbb{1} + \mathbb{1} Y Y \mathbb{1}), \\ W^{B \prec A} &:= \mathbb{1}^\circ + \frac{1}{4}(Z \mathbb{1} X Z), \\ W &:= q W^{A \prec B} + (1 - q + \epsilon) W^{B \prec A} - \epsilon \mathbb{1}^\circ. \end{aligned} \quad (10)$$

Here $\mathbb{1}, X, Y, Z$ are the Pauli matrices and the tensor products between the Hilbert spaces A_I, A_O, B_I, B_O are implicit, as in the remainder of the paper. The process matrix (10) is positive semidefinite for $\epsilon \leq q - 1 + \sqrt{\frac{(1-q)(q+3)}{3}}$ and causally nonseparable for $\epsilon > 0$. As shown in Appendix A, its random robustness is $R_r(W) = \epsilon$. It is maximal for $q = \sqrt{3} - 1 \approx 0.732$, where $\epsilon = \frac{4}{\sqrt{3}} - 2 \approx 0.309$.

The proof that the process (10) cannot be used to violate any causal inequalities, for any local strategy³ consists of two steps: (i) we show that the set of correlations compatible with W is the same as the set of correlations achievable with W^{T_B} (where T_B denotes the partial transpose of the systems $B_I B_O$ with respect to the computational basis); (ii) we verify that W^{T_B} is valid and causally separable, hence cannot violate causal inequalities. Taken together, this establishes that W cannot violate any causal inequalities either and therefore admits a causal model.

² When Alice is given the input x and outputs a (Bob is given an input y and outputs b), no-signaling from Bob to Alice implies that the marginal probability on Alice’s side does not depend on Bob’s input: $\sum_b p_{A \prec B}(ab|xy) = \sum_b p_{A \prec B}(ab|xy'), \forall y, y'$.

³ Note that our proof guarantees the existence of a causal model, which means that the correlations belong to *every* causal polytope, without restriction on the number of inputs and outputs for each party.

The first part of the proof is simple. Using definition (2) and the self-duality of transposition, we rewrite the probability distribution:

$$p(ab|xy) = \text{tr}[W \xi_x^a \otimes \eta_y^b] = \text{tr}[W^{T_B} \xi_x^a \otimes (\eta_y^b)^T]. \quad (11)$$

Additionally, for any quantum instrument [23] $\{\eta_y^b\}$, the instrument $\{\eta_y^{bT}\}$ is also valid, since transposition maps completely positive maps to completely positive maps and trace-preserving maps to trace-preserving maps⁴. This establishes (i), namely that the correlations achievable with W^{T_B} are the same as those compatible with W —note that this holds for *any* process, even when W^{T_B} is *not positive semidefinite*, and therefore not a valid process matrix. In such a case, the probability distribution will be well-defined for local measurements, but not when extending the process with an entangled state between Alice and Bob, an extension which is physically meaningful and to which we will come back in Sec. VI.

For the class of process matrices given in Eq. (10), W^{T_B} is always positive semidefinite. We will now explicitly decompose W^{T_B} as a convex combination of causally ordered process matrices, proving that it is causally separable (formally, this implies that $R_r(W^{T_B}) \leq 0$).

First, one should notice the similarity of $W^{A \prec B}$ with the process matrix $D_{2/3}^{A \prec B}$ of a depolarizing channel (with $\frac{2}{3}$ probability of depolarizing and $\frac{1}{3}$ probability of perfectly transmitting the state) from Alice to Bob

$$D_{2/3}^{A \prec B} := \mathbb{1}^\circ + \frac{1}{12}(\mathbb{1}ZZ\mathbb{1} + \mathbb{1}XX\mathbb{1} - \mathbb{1}YY\mathbb{1}), \quad (12)$$

where only the sign of the term $\mathbb{1}YY\mathbb{1}$ differs compared to $W^{A \prec B}$ of (10). This exactly corresponds to a partial transpose of the systems $B_I B_O$, such that $W^{A \prec B} = (D_{2/3}^{A \prec B})^{T_B}$. Using the definition of the depolarizing process $D_{2/3}^{A \prec B} = \frac{2}{3}\mathbb{1}^\circ + \frac{1}{3}I^{A \prec B}$, where $I^{A \prec B} = \mathbb{1}^{A_I} |I\rangle\langle\langle I|^{A_O B_I} \mathbb{1}^{B_O} / 2$. Since $(W^{B \prec A})^{T_B} = W^{B \prec A}$, we can write W^{T_B} as:

$$\begin{aligned} W^{T_B} &= \frac{2q}{3}\mathbb{1}^\circ + \frac{q}{3}I^{A \prec B} + (1-q+\epsilon)W^{B \prec A} - \epsilon\mathbb{1}^\circ \\ &= \frac{q}{3}I^{A \prec B} + (1-q+\epsilon)W^{B \prec A} + \left(\frac{2q}{3} - \epsilon\right)\mathbb{1}^\circ, \end{aligned} \quad (13)$$

which is a convex decomposition into causally ordered processes as long as $\epsilon \leq \epsilon_c(q) = \frac{2q}{3}$. Since $\epsilon \leq q - 1 + \sqrt{\frac{(1-q)(3+q)}{3}}$ for the process given in Eq. (10) to be valid, and $q - 1 + \sqrt{\frac{(1-q)(3+q)}{3}} \leq \epsilon_c(q)$, the whole class of processes defined in Eq. (10) cannot violate causal inequalities. For a graphical representation of this relationship, see Fig. 2.

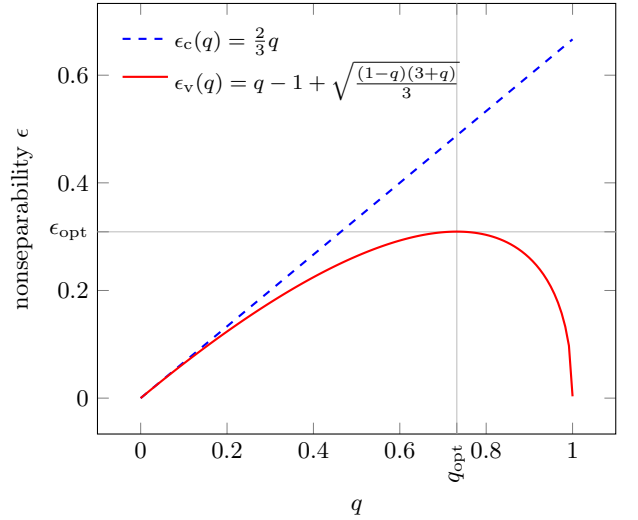


FIG. 2. For W of Eq. (10) to be a valid process, $\epsilon \leq \epsilon_v(q)$ (in the region below the red curve). A process only generates causal correlations for $\epsilon \leq \epsilon_c(q)$ (the region below the dotted blue curve). Since $\epsilon_v \leq \epsilon_c$, every valid process of the form (10) allows for a “causal” probability distribution. ϵ_{opt} and q_{opt} are the parameters maximizing causal nonseparability.

This concludes the proof and provides an explicit causal model for the process W with the instruments $\{\xi_x^a\}$ and $\{\eta_y^b\}$: the process W^{T_B} with the instruments $\{\xi_x^a\}$ and $\{\eta_y^{bT}\}$. Since W^{T_B} is causally separable, it can be interpreted as a probabilistic mixture of two causally ordered processes. There are infinitely many such decompositions since the term $\mathbb{1}^\circ$ in Eq. (13) can be split and added to $\frac{q}{3}I^{A \prec B}$ and $(1-q+\epsilon)W^{B \prec A}$ in any proportion.

W can be taken to be W^{T_B} composed with a transpose map on B . The process matrix $\frac{q}{3}(I^{A \prec B})^{T_B}$ becomes positive (meaning that the associated map is completely positive) when adding at least $\frac{2q}{3}$ of white noise. However, the maximal noise that can be admixed by transferring the $\mathbb{1}^\circ$ term in (13) is $\frac{2q}{3} - \epsilon$, which is strictly smaller. Therefore, the causal model for W suggests a natural interpretation for it as a convex combination of an *unphysical* channel from Alice to Bob (as it is not completely positive) with a *physical* channel from Bob to Alice. This provides some evidence—yet not a proof—that the process W is not physically implementable.

IV. RANDOM CAUSALLY NONSEPARABLE PROCESSES WITH A CAUSAL MODEL

In this section, we develop a method to construct a broad class of causally nonseparable processes with a causal model. Given a random causally separable process, one applies a positive, but not necessarily completely positive map $Q_B(\cdot)$ on Bob’s side:

$$W_{\text{sep}}? = Q_B(W_{\text{sep}}). \quad (14)$$

⁴ The condition on the CJ representation of a CPTP map is that $\text{tr}_{B_O} M^{B_I B_O} = \mathbb{1}^{B_I}$ and it implies $\text{tr}_{B_O} (M^{B_I B_O})^T = \mathbb{1}^{B_I}$.

If the resulting process has negative eigenvalues, it is discarded; otherwise, it is a valid process matrix. Using the same argument as in the preceding section, we know that the resulting process $W_{\text{sep?}}$ will have a causal model. Sometimes—and these are the interesting cases—the process will also be causally nonseparable. This can readily be checked this via SDP (7).

We generated causally separable process matrices W_{sep} (where $d_{A_I} = d_{B_I} = d_{A_O} = d_{B_O} = 2$) according to an asymptotically uniform distribution using the “hit-and-run” technique (see Appendix C for details) and Eq. (14), using the transposition map ${}^{\text{T}_B}$ for Q_B . We found that most (69%) of the resulting matrices were positive and hence valid process matrices. About half of these turn out to be causally nonseparable while—by construction—allowing for a causal model (we denote this set by $\mathcal{W}_{\text{nsep}}^{(c)}$). The histogram of the resulting causal nonseparabilities is shown in Fig. 3. There is therefore a *finite* probabil-

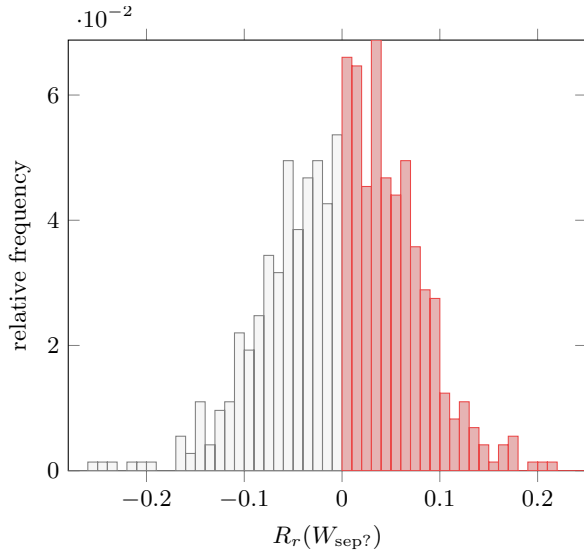


FIG. 3. Histogram of the random robustness R_r of the subset of 690 ($\approx 69\%$) valid processes $W_{\text{sep?}}$ generated from 1000 uniformly distributed causally separable processes (see Appendix C) and applying partial transposition on Bob’s side. The 366 ($\approx 53\%$) which are causally nonseparable, while admitting a causal model, are represented in red.

ity of generating a nonseparable process with a causal model starting from a random of causally separable process. Since the map ${}^{\text{T}_B}$ is measure-preserving, the set of causally nonseparable processes which admit a causal model is of the same dimension as the set of valid processes \mathcal{W} itself (see Appendix B for details).

V. “WERNER” CAUSALLY NONSEPARABLE PROCESSES

We will denote the set of causally nonseparable processes as $\mathcal{W}_{\text{nsep}}$. It is composed of the set of processes

with a causal model ($\mathcal{W}_{\text{nsep}}^{(c)}$) and the set of processes that can violate causal inequalities ($\mathcal{W}_{\text{nsep}}^{(\text{nc})}$), which are graphically represented in Fig. 4.

We now construct bipartite processes violating a causal inequality, but which, mixed with some amount of “white noise” $\mathbb{1}^\circ$, turn into causally nonseparable processes with a causal model. This behavior is reminiscent of “Werner states”, which violate Bell inequalities until a noise level of up to $\frac{1}{2}$ but are entangled when mixed with noise up to a level of $\frac{2}{3}$ [9]. It shows that the analogy between causal inequalities and causal nonseparability, on the one hand, and Bell inequalities and entanglement, on the other hand, applies to the two-party case.

The idea is to use a convex combination of a process in $\mathcal{W}_{\text{nsep}}^{(c)}$ and a process in $\mathcal{W}_{\text{nsep}}^{(\text{nc})}$, which is also invariant under partial transposition with respect to B ⁵. In this way, one can generate a broad class of “Werner causally nonseparable processes”.

We will use the process defined in Eq. (10) with the maximal causal nonseparability:

$$W_{\text{opt}} := (\sqrt{3} - 1)W^{A \prec B} + \frac{1}{\sqrt{3}}W^{B \prec A} - \left(\frac{4}{\sqrt{3}} - 2\right)\mathbb{1}^\circ, \quad (15)$$

together with the process

$$W_{\text{OCB}} := \mathbb{1}^\circ + \frac{1}{4\sqrt{2}}(\mathbb{1}ZZ\mathbb{1} + Z\mathbb{1}XZ), \quad (16)$$

which was proposed and shown to violate causal inequalities in Ref. [3]. In Appendix A, we show that the resulting mixture

$$W_{\text{mix}}(\alpha) := \alpha W_{\text{opt}} + (1 - \alpha)W_{\text{OCB}} \quad (17)$$

has nonseparability $R_{\text{mix}}(\alpha) := R_r(W_{\text{mix}}(\alpha)) = \alpha R_r(W_{\text{opt}}) + (1 - \alpha)R_r(W_{\text{OCB}})$, where $R_r(W_{\text{OCB}}) = \sqrt{2} - 1$ is the random robustness of W_{OCB} .

Following the same argument as in the previous section, we can now examine the causal nonseparability of $W_{\text{mix}}^{\text{T}_B}$, since, by transferring the partial transpose onto Bob’s CP maps, we know that it can produce exactly the same correlations as W_{mix} . Its nonseparability $R'_{\text{mix}}(\alpha) := R_r(W_{\text{mix}}^{\text{T}_B}(\alpha))$ is again the weighted average (see Appendix A) of the nonseparabilities of $W_{\text{opt}}^{\text{T}_B}$ and of $W_{\text{OCB}}^{\text{T}_B} = W_{\text{OCB}}$:

$$\begin{aligned} R'_{\text{mix}}(\alpha) &= \alpha R_r(W_{\text{opt}}^{\text{T}_B}) + (1 - \alpha)R_r(W_{\text{OCB}}) \\ &= R_{\text{mix}}(\alpha) + \frac{2\alpha}{3}R_r(W_{\text{opt}}^{\text{T}_B}) < R_{\text{mix}}(\alpha), \end{aligned} \quad (18)$$

⁵ Or, in the more general scenario described in the previous section, is invariant under the non-completely-positive operation $Q_B(\cdot)$.

where we used $R_r(W_{\text{opt}}^{\text{T}_B}) = \frac{2\sqrt{3}-4}{3} < 0$ and $R_r(W_{\text{OCB}}) = \sqrt{2} - 1$. This means that there is a finite gap between the nonseparability of W_{mix} and the nonseparability of $W_{\text{mix}}^{\text{T}_B}$.

Using a see-saw algorithm [4], we numerically verified that $W_{\text{mix}}(\alpha)$ indeed violates the causal inequalities of Ref. [4] as long as $W_{\text{mix}}^{\text{T}_B}$ is nonseparable ($R'_{\text{mix}}(\alpha) > 0$), i.e., when $\alpha < \frac{3(\sqrt{2}-1)}{1+3\sqrt{2}-2\sqrt{3}} \approx 0.6987$.

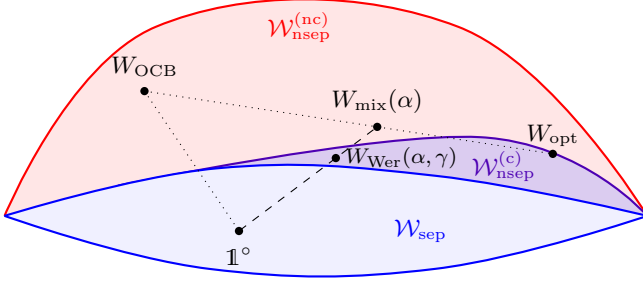


FIG. 4. Schematic depiction of different types of process matrices. $W_{\text{nsep}}^{(\text{nc})}$ is the set of process matrices which can violate causal inequalities and $W_{\text{nsep}}^{(\text{c})}$ is the set of nonseparable process matrices which admit a causal model and W_{sep} is the set of separable processes. The process $W_{\text{mix}} := \alpha W_{\text{opt}} + (1 - \alpha)W_{\text{OCB}}$, where W_{opt} admits a causal model and W_{OCB} doesn't, gives rise to “Werner type” processes $W_{\text{Wer}}(\gamma, \alpha) := (1 - \gamma)W_{\text{mix}}(\alpha) + \gamma 1^\circ$, which have a causal model but are causally nonseparable for a certain level of noise (21).

This gap translates into a gap between the level of white noise 1° that W_{mix} can tolerate before admitting a causal model and the level of noise at which it becomes nonseparable. We therefore define the “Werner process” as a convex combination of 1° and $W_{\text{mix}}(\alpha)$:

$$W_{\text{Wer}}(\gamma, \alpha) := (1 - \gamma)W_{\text{mix}}(\alpha) + \gamma 1^\circ. \quad (19)$$

Using the definition of nonseparability (7), one can verify that the following relations hold (see Appendix A):

$$\begin{aligned} W_{\text{Wer}}\left(\gamma < \frac{R_{\text{mix}}(\alpha)}{1 + R_{\text{mix}}(\alpha)}, \alpha\right) &\in W_{\text{nsep}}, \\ W_{\text{Wer}}^{\text{T}_B}\left(\gamma \geq \frac{R'_{\text{mix}}(\alpha)}{1 + R'_{\text{mix}}(\alpha)}, \alpha\right) &\in W_{\text{sep}}. \end{aligned} \quad (20)$$

As W_{Wer} can violate causal inequalities only if $W_{\text{Wer}}^{\text{T}_B}$ is causally nonseparable (remember the proof of Sec. III), we conclude from (20) that

$$W_{\text{Wer}}\left(\frac{R'_{\text{mix}}(\alpha)}{1 + R'_{\text{mix}}(\alpha)} \leq \gamma < \frac{R_{\text{mix}}(\alpha)}{1 + R_{\text{mix}}(\alpha)}, \alpha\right) \in W_{\text{nsep}}^{(\text{c})}, \quad (21)$$

which mimics the behavior of Werner states. See Fig. 4 for a graphical representation of the location of W_{Wer} with respect to the different sets of processes.

VI. RELATIONSHIP TO EXTENSIBLY CAUSAL PROCESSES

In the context of the physical implementability of process matrices, it is natural to consider the *extension* of a process matrix with an *entangled state shared between the parties*. A process is “extensibly causal” if, even when extended with a shared entangled state, it *cannot violate causal inequalities* [8].

Extending a physically implementable process with an entangled state shared among the parties should also result in a physically implementable process. On this account, it is important to consider not only whether a process has a causal model, but rather whether it has such a model when extended with an entangled state. In Ref. [8], an example of a tripartite process with a causal model but which is not extensibly causal was presented, showing that both notions really differ and that the violation of causal inequalities can be “activated” by entanglement—and for which no physical implementation is known.

Note that the proof (Sec. III) of the existence of causal model for W does not hold when the process is extended with an entangled state between Alice and Bob. It therefore cannot prove that W is extensibly causal. It crucially relies on the fact that the transpose $\{(\eta_y^b)^T\}$ of a valid instrument for Bob $\{\eta_y^b\}$ is still a valid instrument. However, taking the full transpose on Bob’s instrument would lead to a “causal model” with a partial transpose of the shared entangled state, which can lead to negative probabilities. Conversely, the *partial transpose of Bob’s instrument* (with no transposition on Bob’s part of the entangled state) is not a valid instrument and does not yield positive probabilities in general.

To numerically study whether W_{opt} from Eq. (15) is extensibly causal, we extended it with a maximally entangled state of two ququarts ($|\phi\rangle^{A_I B_I'} := \frac{1}{2}(|00\rangle + |11\rangle + |22\rangle + |33\rangle)$):

$$W_{\text{ext}} := W_{\text{opt}} \otimes |\phi\rangle \langle \phi|^{A_I B_I'}. \quad (22)$$

We chose a maximally entangled ququart state because we believe that extending W_{opt} with a higher dimensional state would not improve its ability to violate causal inequalities.

Using the see-saw algorithm, we optimized W_{ext} for a violation of the simplest causal inequalities [4]. We found that W_{ext} is able to violate (by about $8 \cdot 10^{-5}$) the GYNI inequality (9), which proves that W_{opt} is *not* extensibly causal. Incidentally, it also shows that *in the bipartite case as well*, the violation of causal inequalities can be “activated” using entanglement.

If we adopt the view that extensibly causal processes are physical, the activation of violation of causal inequalities, this suffices to exclude W_{opt} in the same way as the tripartite process with a causal model but which is not extensibly causal, given in Ref. [8], independently of the argument based on the decomposition of Sec. III. How-

ever, this is not possible anymore when admixing a small amount of white noise: $(1 - \kappa)W_{\text{ext}} + \kappa\mathbb{1}^\circ$. We ran the see-saw algorithm for different levels of white noise and found no violation of GYNI for $\kappa > 3.3 \cdot 10^{-4}$ (see Fig. 5 for a graphical representation of the relationship between noise and violation of GYNI). Similarly, neither the other causal inequality from Ref. [4] nor the “original” causal inequality from Ref. [3] could be violated through see-saw optimization.

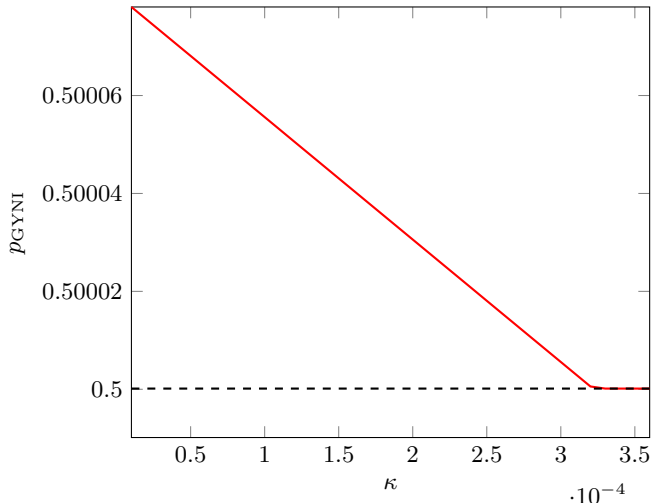


FIG. 5. Numerically optimized (see-saw) violation of the GYNI inequality Eq. (9), using the noisy extended process $(1 - \kappa)W_{\text{ext}} + \kappa\mathbb{1}^\circ$ (see Eq. (22)) and causal bound (dashed). For a noise level of $\kappa > 3.3 \cdot 10^{-4}$, the algorithm fails to find a strategy violating the inequality, as well as for the other known bipartite causal inequalities [3, 4].

This gives reasonable evidence⁶ that W_{opt} mixed with very little white noise is extensively causal, while still being causally nonseparable (see Fig. 6 for a graphical representation). The argument for unphysicality of Sec. III still applies to it. This leads us to conjecture that some extensively causal processes cannot be physically realized.

VII. CONCLUSIONS

We studied the classification of causally nonseparable process matrices for two parties and found that composing a class of causally separable processes with a transpose map on one party’s side results in nonseparable processes with causal models, i.e., that cannot violate causal inequalities.

⁶ It falls short of being a proof because (i) the entangled state added in W_{ext} is finite-dimensional; (ii) the see-saw technique is not guaranteed to converge to the global optimum; (iii) only the three *known* bipartite causal inequalities were tested, inequalities with more settings might still be violated.

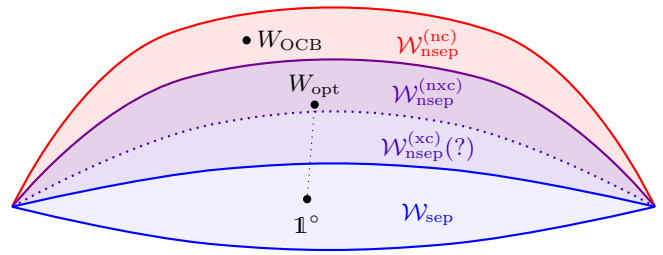


FIG. 6. Schematic depiction of the sets process matrices with respect to extendible causal separability. W_{sep} is the set of separable processes, $W_{\text{nsep}}^{(\text{nc})}$ is the set of process matrices which can violate causal inequalities, $W_{\text{nsep}}^{(\text{nxc})}$ the set of processes with a causal model but which can violate causal inequalities when extended with entanglement, such as W_{opt} . Based on our numerical evidence (see Fig. 5), we conjecture that the set $W_{\text{nsep}}^{(\text{xc})}$ of processes which are causally nonseparable and “extensively causal” is not empty. Note that the set of nonseparable processes with a causal model is $W_{\text{nsep}}^{(\text{c})} = W_{\text{nsep}}^{(\text{xc})} \cup W_{\text{nsep}}^{(\text{nxc})}$.

Since the only interpretation we know relies on applying a non-completely positive map (which is itself is unphysical) to a valid process, the conjecture that processes which do not violate causal inequalities are physically implementable is undermined.

We also provided a simple algorithm to generate non-separable processes with causal models—starting from a random separable process and composing it with a positive, but not completely positive map on one party’s side. With a finite probability, this yields a nonseparable process with a causal model and shows that the measure of such processes is nonzero within the space of valid processes. The “hit-and-run” algorithm we used to generate random process matrices might be of independent interest.

We then developed the analogy between entanglement/nonlocality and causal nonseparability/noncausal correlations by providing a process analogous to a Werner state: it starts having a causal model when mixed with a certain amount of white noise, while still being strictly causally nonseparable.

Finally, we studied whether our processes still have a causal model when extended with an entangled state shared between Alice and Bob (whether they are “extensively causal”). The numerical evidence prompted us to conjecture that some of the nonseparable processes we studied are extensively causal, while not being physically implementable.

An important question remains open: if some processes which have an (extensible) causal model are nonphysical, which other criterion should be used to rule them out? One fairly natural approach is to postulate a “purification principle”, according to which physically realizable processes can be recovered as part of a pure process in a larger space [24].

ACKNOWLEDGMENTS

We thank Ognian Oreshkov and Fabio Costa for pointing out the significance of “extensible causal separability” in the context of physical implementability, and Ralph

Silva for a useful comment on an earlier version of this manuscript. We acknowledge support from the European Commission project RAQUEL (No. 323970); the Austrian Science Fund (FWF) through the Special Research Programme FoQuS, the Doctoral Programme CoQuS and Individual Project (No. 2462) and the John Templeton Foundation.

-
- [1] J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics* **1**, 195–200 (1964).
 - [2] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed Experiment to Test Local Hidden-Variable Theories,” *Phys. Rev. Lett.* **23**, 880–884 (1969).
 - [3] O. Oreshkov, F. Costa, and Č. Brukner, “Quantum correlations with no causal order,” *Nat. Commun.* **3**, 1092 (2012).
 - [4] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, “The simplest causal inequalities and their violation,” *New J. Phys.* **18**, 013008 (2016).
 - [5] Ä. Baumeler, A. Feix, and S. Wolf, “Maximal incompatibility of locally classical behavior and global causal order in multiparty scenarios,” *Phys. Rev. A* **90**, 042106 (2014).
 - [6] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, “Quantum computations without definite causal structure,” *Phys. Rev. A* **88**, 022318 (2013).
 - [7] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, “Witnessing causal nonseparability,” *New J. Phys.* **17**, 102001 (2015).
 - [8] O. Oreshkov and C. Giarmatzi, “Causal and causally separable processes,” [arXiv:1506.05449](https://arxiv.org/abs/1506.05449).
 - [9] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” *Phys. Rev. A* **40**, 4277–4281 (1989).
 - [10] Č. Brukner, “Quantum causality,” *Nat. Phys.* **10**, 259–263 (2014).
 - [11] R. L. Smith, “Monte Carlo procedures for generating random feasible solutions to mathematical programs,” in *A Bulletin of the ORSA/TIMS Joint National Meeting, Washington, DC*, vol. 101. 1980.
 - [12] R. L. Smith, “Efficient Monte Carlo Procedures for Generating Points Uniformly Distributed over Bounded Regions,” *Operations Research* **32**, 1296–1308 (1984).
 - [13] Č. Brukner, “Bounding quantum correlations with indefinite causal order,” *New J. Phys.* **17**, 083034 (2015).
 - [14] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Theoretical framework for quantum networks,” *Phys. Rev. A* **80**, 022339 (2009).
 - [15] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra Appl.* **10**, 285–290 (1975).
 - [16] A. Jamiołkowski, “Linear transformations which preserve trace and positive semidefiniteness of operators,” *Rep. Math. Phys.* **3**, 275–278 (1972).
 - [17] G. Gutoski and J. Watrous, “Toward a General Theory of Quantum Games,” in *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing, STOC ’07*, pp. 565–574. ACM, New York, NY, USA, 2007.
 - [18] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Transforming quantum operations: Quantum supermaps,” *EPL* **83**, 30004 (2008).
 - [19] M. S. Leifer and R. W. Spekkens, “Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference,” *Phys. Rev. A* **88**, 052130 (2013).
 - [20] D. Chruściński and G. Sarbicki, “Entanglement witnesses: construction, analysis and classification,” *J. Phys. A: Math. Theor.* **47**, 483001 (2014).
 - [21] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, “Device-Independent Security of Quantum Cryptography against Collective Attacks,” *Phys. Rev. Lett.* **98**, 230501 (2007).
 - [22] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, “Experimental superposition of orders of quantum gates,” *Nat. Commun.* **6**, 7913 (2015).
 - [23] E. B. Davies and J. T. Lewis, “An operational approach to quantum probability,” *Comm. Math. Phys.* **17**, 239–260 (1970).
 - [24] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, “A purification postulate for quantum mechanics with no causal order,” (*in preparation*).
 - [25] L. Lovász, “Hit-and-run mixes fast,” *Math. Program.* **86**, 443–461 (1999).

Appendix A: Analytic proofs for nonseparabilities in the main text

First, we prove that the class of processes, defined by

$$\begin{aligned}
 W^{A \prec B} &:= \mathbb{1}^\circ + \frac{1}{12}(\mathbb{1}ZZ\mathbb{1} + \mathbb{1}XX\mathbb{1} + \mathbb{1}YY\mathbb{1}), \\
 W^{B \prec A} &:= \mathbb{1}^\circ + \frac{1}{4}(Z\mathbb{1}XZ), \\
 W &:= qW^{A \prec B} + (1 - q + \epsilon)W^{B \prec A} - \epsilon\mathbb{1}^\circ, \quad (A1)
 \end{aligned}$$

has causal separability $R_r(W) = \epsilon$. To do so, we define a causal witness S_W for it:

$$S_W = \mathbb{1}^\circ - \frac{1}{4}(\mathbb{1}ZZ\mathbb{1} + \mathbb{1}XX\mathbb{1} + \mathbb{1}YY\mathbb{1}) - \frac{1}{4}(Z\mathbb{1}XZ). \quad (A2)$$

We first verify that S_W is a causal witness: $\text{tr}_{A_O} S_W \geq 0$ and $\text{tr}_{B_O} S_W \geq 0$, which is a sufficient condition [7] for S to have positive trace with any causally separable

process. Therefore, $\text{tr}[W_{\text{sep}}S_W] \geq 0$ and S_W is indeed a causal witness.

We compute $\text{tr}[S_W(W + \lambda \mathbb{1}^\circ)] = -\epsilon + \lambda$, which is negative for $\lambda < \epsilon$ and implies that $R_r(W) \geq \epsilon$. From Eq. (A1), it is clear that $W + \lambda \mathbb{1}^\circ$ is causally separable for $\lambda \geq \epsilon$, so $R_r(W) \leq \epsilon$. This establishes that $R_r(W) = \epsilon$.

Using the same approach, we can show that the process

$$W_{\text{mix}}(\alpha) := \alpha W_{\text{opt}} + (1 - \alpha)W_{\text{OCB}}, \quad (\text{A3})$$

where W_{opt} is defined in (15) and W_{OCB} in (16), has nonseparability $R_r(W_{\text{mix}}(\alpha)) = \alpha R_r(W_{\text{opt}}) + (1 - \alpha)R_r(W_{\text{OCB}}) = 1 + \alpha(\frac{4}{\sqrt{3}} - 3)$.

The convexity of random robustness [7] implies that the random robustness of a convex combination is smaller than the convex combination of the random robustnesses, so $R_r(W_{\text{mix}}) \leq 1 + \alpha(\frac{4}{\sqrt{3}} - 3)$. Using the same witness S_W as before, we compute $\text{tr}[S_W(W_{\text{mix}} + \lambda \mathbb{1}^\circ)] = -1 - \alpha(\frac{4}{\sqrt{3}} - 3) + \lambda$ which is strictly negative when $\lambda > 1 + \alpha(\frac{4}{\sqrt{3}} - 3)$ and implies that $R_r(W_{\text{mix}}) \geq 1 + \alpha(\frac{4}{\sqrt{3}} - 3)$. We conclude that $R_r(W_{\text{mix}}) = 1 + \alpha(\frac{4}{\sqrt{3}} - 3)$.

The same proof (with the same witness S_W given in Eq. (A2)) can also be used to show that the causal nonseparability of $W_{\text{mix}}^{\text{TB}}$ is again the convex combination $R_r(W_{\text{mix}}^{\text{TB}}) = \alpha R_r(W_{\text{opt}}^{\text{TB}}) + (1 - \alpha)R_r(W_{\text{OCB}})$.

Appendix B: The dimension of the set causally separable processes

Here we show that the set of causally separable processes \mathcal{W}_{sep} has the same dimension as the set of valid processes \mathcal{W} , which establishes that the set of causally separable processes has nonzero measure in the set of valid processes.

We will use the Hilbert-Schmidt decomposition of operators. An arbitrary process W can be decomposed as $W = \sum_{ijkl=0}^3 \alpha_{ijkl} \sigma_i^{A_I} \otimes \sigma_j^{A_O} \otimes \sigma_k^{B_I} \otimes \sigma_l^{B_O}$.

The condition of normalization of probabilities, i.e., $\text{tr}[W \cdot M_A^{\text{CPTP}} \otimes M_B^{\text{CPTP}}] = 1$ for all CJ-representations of completely positive *trace-preserving maps* M_A^{CPTP} and M_B^{CPTP} , implies that some terms of the Hilbert-Schmidt decomposition, corresponding to “causal loops” are excluded. In particular, $\alpha_{0j00} = \alpha_{000l} = \alpha_{0j0l} = \alpha_{0jkl} = \alpha_{ij0l} = \alpha_{ijkl} = 0$ for $i, j, k, l \geq 1$ (see the Supplementary Material of Ref. [3]).

Counting all the “allowed terms” in the Hilbert-Schmidt decomposition, we find that the dimension d_W of \mathcal{W} is:

$$d_W = (1 + d_{A_I}^2(d_{A_O}^2 - 1))(d_{B_I}^2 - 1) + (d_{A_I}^2 - 1)d_{B_I}^2 d_{B_O}^2.$$

For causally ordered processes in $W^{A \prec B} \in \mathcal{W}^{A \prec B}$ compatible with the causal order $A \prec B$, some additional terms, which allow for signaling from Bob to Alice, are excluded in the Hilbert-Schmidt decomposition, reducing the dimension to

$$d_{W^{A \prec B}} = d_{A_I}^2(1 + (d_{B_I}^2 - 1)d_{A_O}^2) - 1.$$

This means that the set of causally ordered processes $\mathcal{W}^{A \prec B}$ has measure zero within the set of all process matrices.

Separable processes are convex combinations of $W^{A \prec B}$ and $W^{B \prec A}$. This means that all the terms allowed in the Hilbert-Schmidt decomposition of a valid process matrix are also allowed in the decomposition of separable processes. Therefore \mathcal{W} and \mathcal{W}_{sep} share the same basis and $d_W = d_{W_{\text{sep}}}$.

Appendix C: Generating uniformly distributed processes

We consider the space of process matrices \mathcal{W} as being embedded in \mathbb{R}^{d_W} . We wish to obtain a *uniform sample* of \mathcal{W} according to the d_W -dimensional volume (Lebesgue measure), which also corresponds to the measure generated by the Hilbert-Schmidt metric. We use an adaptation of the “hit-and-run” Markov chain sampler [11, 12] for this task. The iteration works as follows:

Algorithm 1

1. Select a starting point W_0 .
2. Choose a traceless matrix Q_{i+1} from a set of d_W orthogonal traceless matrices and generate a random sign variable $s = \pm 1$.
3. Find μ such that $W_i + \mu(\mathbb{1}^\circ + sQ_{i+1})$ is on the boundary of the set of valid processes.
4. Generate a random real scalar $\theta \in [0, \mu]$. Take $W_{i+1} = W_i + \theta(\mathbb{1}^\circ + sQ_{i+1})$ and go to step 2.

The set of directions is simply the Hilbert-Schmidt basis of allowed terms; it has dimension d_W (see Appendix B). For bipartite processes with $d_{A_I} = d_{B_I} = d_{A_O} = d_{B_O} = 2$, there are $d_W = 87$ possible directions to choose from.

Finding the intersection with the boundary of the set of positive processes in step 3 turns out to be a semidefinite program

$$\begin{aligned} & \max \mu \\ & \text{s.t. } W_i + \mu(\mathbb{1}^\circ + sQ_{i+1}) \geq 0. \end{aligned} \quad (\text{C1})$$

However, the SDP which computes μ at each step of the Markov chain is a bottleneck of the algorithm. Instead, we can skip it and generate $\theta \in [0, 1]$, rejecting and retrying if the resulting process is not positive:

Algorithm 2

1. Select a starting point W_0 .
2. Choose a traceless matrix Q_{i+1} from a set of d_W orthogonal traceless matrices and generate a random sign variable $s = \pm 1$.
3. Generate a random real scalar $\theta \in [0, 1]$. Take $W_{i+1} = W_i + \theta(\mathbb{1}^\circ + sQ_{i+1})$.

4. If $W_{i+1} \geq 0$, go to step 2, otherwise repeat step 3.

The matrices $\mathbb{1}^\circ + Q_i$ are chosen to be slightly *outside* the set \mathcal{W} by having slightly negative eigenvalues. Therefore, there is always a finite probability of rejection at step 4, which guarantees that the algorithm samples uniformly all the way up to the boundary.

The resulting sample $\{W_i\}_{i=0}^\infty$ is uniform when two conditions hold [12]. First, from every W_i, W' the probability to have $W_{i+d_W} = W'$ is nonzero, which is indeed true: In d_W steps, one can reach any W' starting from any W_i . Second, the uniform distribution is a stationary distribution of the Markov chain. This is also the case: for any W_i, W' the probability to reach W' starting from W_i in d_W steps is the same as the probability to reach W_i starting from W' in d_W steps.

An upper bound on the *convergence* of the hit-and-run algorithm for convex sets (which is the case for the set \mathcal{W}) is known—in particular, the mixing time scales as $\tilde{O}(d_w^3) := O(d_w^3 \text{polylog } d_w)$, which matches the best known mixing times for other algorithms [25]. For $d_w = 87$ (which is the case for $d_{A_I} = d_{B_I} = d_{A_O} = d_{B_O} = 2$), we would need around $7 \cdot 10^7$ samples to achieve the same statistical significance as from a one-dimensional hit-and-run with 100 samples, which we deem sufficient for our purposes.

To sample uniformly distributed *causally separable* processes, we use the *rejection method*: after sampling

$7 \cdot 10^7$ process matrices (with a warm-up period of 10^6 discarded steps), we randomly select 1000 causally separable processes (rejecting the $\approx 92.5\%$ of nonseparable ones) in the sample.

The map ${}^{\text{T}_B}$ preserves the Lebesgue measure, since it corresponds to reflections in \mathbb{R}^{d_W} . Therefore, if upon applying ${}^{\text{T}_B}$ to random causally separable matrices, there is a finite probability to obtain a *valid, causally nonseparable process*, this means that the set of causally nonseparable processes with a separable partial transpose is full dimensional (see Fig. 7).

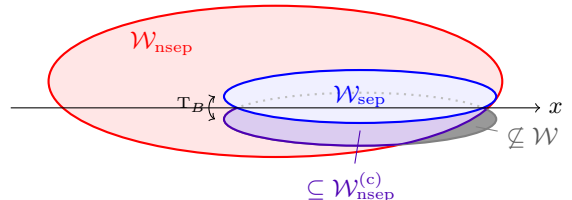


FIG. 7. Schematic two-dimensional cut of \mathcal{W} in \mathbb{R}^{d_W} , the partial transpose ${}^{\text{T}_B}$ here corresponds to a reflection along the horizontal axis x . The partial transpose of the set of causally separable processes \mathcal{W}_{sep} consists of three parts: (i) causally separable matrices ($\subseteq \mathcal{W}_{\text{sep}}$), (ii) non-valid processes ($\not\subseteq \mathcal{W}$) and (iii) valid, causally nonseparable matrices with a causal model ($\subseteq \mathcal{W}_{\text{nsep}}^{(c)}$).